

University of Groningen

## Computer-aided detection in breast MRI

Dorrius, Monique D.; Jansen-van der Weide, Marijke C.; van Ooijen, Peter M. A.; Pijnappel, Ruud M.; Oudkerk, Matthijs

*Published in:*  
European Radiology

*DOI:*  
[10.1007/s00330-011-2091-9](https://doi.org/10.1007/s00330-011-2091-9)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2011

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Dorrius, M. D., Jansen-van der Weide, M. C., van Ooijen, P. M. A., Pijnappel, R. M., & Oudkerk, M. (2011). Computer-aided detection in breast MRI: a systematic review and meta-analysis. *European Radiology*, 21(8), 1600-1608. <https://doi.org/10.1007/s00330-011-2091-9>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Computer-aided detection in breast MRI: a systematic review and meta-analysis

Monique D. Dorrius · Marijke C. Jansen-van der Weide · Peter M. A. van Ooijen ·  
Ruud M. Pijnappel · Matthijs Oudkerk

Received: 10 September 2010 / Revised: 1 December 2010 / Accepted: 12 January 2011 / Published online: 15 March 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

## Abstract

**Objectives** To evaluate the additional value of computer-aided detection (CAD) in breast MRI by assessing radiologists' accuracy in discriminating benign from malignant breast lesions.

**Methods** A literature search was performed with inclusion of relevant studies using a commercially available CAD system with automatic colour mapping. Two independent researchers assessed the quality of the studies. The accuracy of the radiologists' performance with and without CAD was presented as pooled sensitivity and specificity.

**Results** Of 587 articles, 10 met the inclusion criteria, all of good methodological quality. Experienced radiologists reached comparable pooled sensitivity and specificity before and after using CAD (sensitivity: without CAD: 89%; 95% CI: 78–94%, with CAD: 89%; 95% CI: 81–94%) (specificity: without CAD: 86%; 95% CI: 79–91%, with CAD: 82%; 95% CI: 76–87%). For residents the pooled sensitivity increased from 72% (95% CI: 62–81%) without CAD to 89% (95% CI: 80–94%) with CAD, however, not significantly. Concerning specificity, the results were similar (without CAD: 79%; 95% CI: 69–86%, with CAD: 78%; 95% CI: 69–84%).

**Conclusions** CAD in breast MRI has little influence on the sensitivity and specificity of experienced radiologists and therefore their interpretation remains essential. However, residents or inexperienced radiologists seem to benefit from CAD concerning breast MRI evaluation.

**Keywords** Magnetic resonance imaging · Breast · Computer aided detection · CAD · Meta-analysis

## Introduction

Dynamic contrast-enhanced Magnetic Resonance Imaging (MRI) is increasingly used to evaluate pathological features of the breast. Applications for MRI of the breast include diagnostic and screening indications [1–6]. Image analysis is based on the enhancement pattern of lesions in dynamic breast MRI and on morphological characteristics [7–9]. Using those two criteria for the interpretation of the images, breast MRI has a very high sensitivity, which usually exceeds 90% [10–12] and a negative breast MRI shows a sufficient high negative predictive value (NPV) (97%) to safely rule out malignancy [13–15]. However, breast MRI has several limitations, the overall reported specificity varies between 67% and 72%, which therefore results in a high number of false-positive results [10, 12, 16]. Furthermore, MRI requires significant time for image acquisition, processing and interpretation [17, 18]. In order to try to overcome those limitations, Computer Aided Detection (CAD) programs for MR imaging of the breast have been developed [18]. In general, CAD software was developed to identify suspect features on the image and bring them to the attention of the radiologist, in order to decrease false-negative readings [19]. However, in breast MRI, most lesions were regarded as having already been detected by the radiologist. Therefore, the primary aim to develop CAD for breast MRI was not to identify lesions, but to assist the radiologist in determining which lesions are benign and which are malignant.

Computer-aided detection systems automate many processing and analysis functions, which would normally have

M. D. Dorrius (✉) · M. C. J.-v. der Weide · P. M. A. van Ooijen ·  
R. M. Pijnappel · M. Oudkerk  
Department of Radiology, Center for Medical Imaging,  
University Medical Center Groningen,  
Hanzeplein 1, PO box 30.001, 9700 RB Groningen,  
the Netherlands  
e-mail: m.d.dorrius@rad.umcg.nl

to be performed manually by MRI technologists and radiologists. The automated kinetic assessment of CAD generates a colour-coding based on the signal intensity voxel changes during the enhancement of the breast tissue. This provides an easier way of interpreting the patterns of contrast enhancement (persistent, plateau and washout enhancement) across a series of images, which may help identify lesions and their likelihood of being malignant.

The implementation of CAD software may improve the accuracy of breast MRI by reducing the number of false-positive diagnoses and by shortening the time needed to interpret breast MRI images [17, 18, 20, 21]. Furthermore, a state of the art CAD system should automatically identify (almost) all non-calcified lesions suspected of malignancy at mammography. This is reflected by a very high sensitivity and NPV for these non-calcified breast lesions.

The purpose of this systematic review and meta-analysis is to assess the radiologists' accuracy in discriminating benign from malignant breast lesions regarding breast MRI with and without CAD implementation in terms of sensitivity and specificity.

## Materials and methods

### Search strategy

A computerised search was performed to identify all relevant studies in Medline and Embase up to 2010. The following search terms were used in Medline: “Diagnosis, Computer-Assisted” [Mesh term] OR “computer-aided-diagnosis” [Text Word] OR “computer-aided-detection” [Text Word] OR “computer-aided” [Text Word] OR “CAD” [Text Word] OR “three-time-point method” [Text Word] AND “Magnetic Resonance Imaging” [Mesh term] OR “Magnetic Resonance Imaging” [Text Word] OR “MRI” [Text Word] OR “MR imaging” [Text Word] AND “Breast Neoplasms” [Mesh term] OR “breast cancer” [Text Word] OR “mamma carcinoma” [Text Word] OR “malignant breast lesions” [Text Word] AND (“Sensitivity and Specificity” [Mesh term] OR “specificity OR sensitivity” [Text Word]). In Embase the same strategy was used. All languages were considered.

### Eligibility criteria and study selection

We searched for studies assessing the value of CAD for a radiologist in the discrimination between benign and malignant breast lesions with MRI. Studies were included if the following inclusion criteria were met: (1) all patients had undergone breast MRI; (2) a commercially available CAD system was used; (3) the study population had benign and malignant breast lesions; (4) accuracy, sensitivity,

specificity, positive predictive value and/or negative predictive value was/were measured or could be derived, and (5) studies had to be published with original data in peer-reviewed journals. Studies in which an institution-specific CAD system was used were excluded, as well as reviews, editorials and case reports.

Two researchers (MDD, MCJW) independently selected relevant studies based on title and abstract or full article. Any discrepancies concerning the study selection were resolved by discussion of the full article. The complete search yielded 587 studies. 519 out of 587 studies were excluded based on the title. After removing duplicates ( $n=12$ ), 56 studies were screened on title and abstract. Twenty-nine studies did not meet the inclusion criteria (technical article ( $n=20$ ), overview ( $n=8$ ) and case report ( $n=1$ )). From the remaining 27 studies the full article was reviewed. Seventeen studies were excluded because CAD was an institution-specific CAD system. Ten studies [20–29] fulfilled our inclusion criteria (Fig. 1). Eight studies were in the English language [20–24, 26, 27, 29]; the other 2 were in the German Language [25, 28].

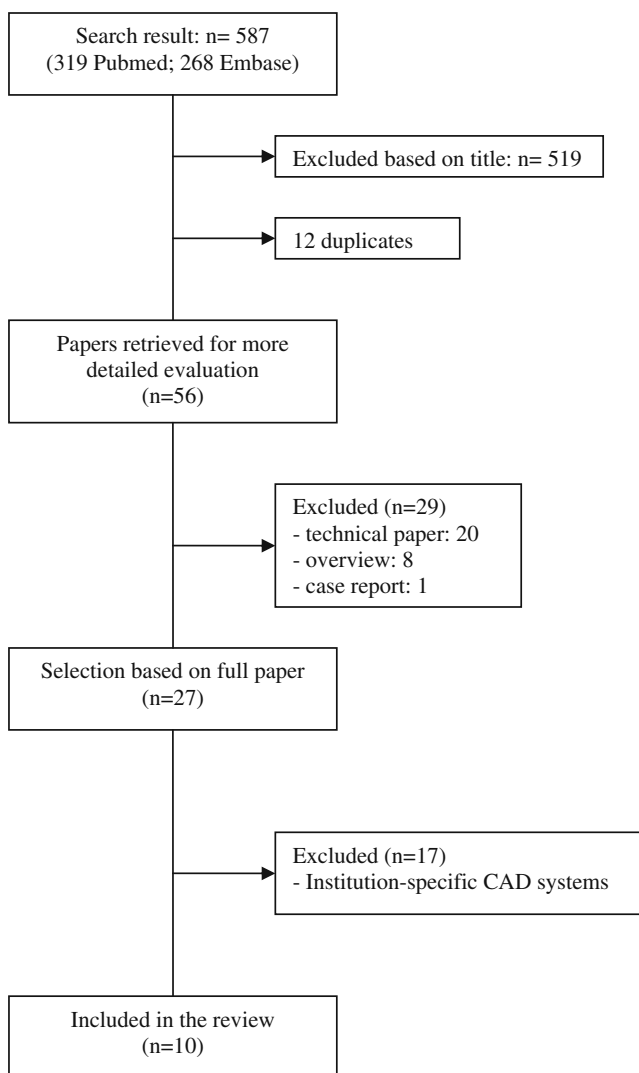
### Data collection and quality assessment

The following study descriptives were extracted: population descriptives (age, number of patients, number of benign and malignant lesions), study design, type of MRI used, type of CAD software used, minimum threshold enhancement used, number of radiologists that assessed the MR images with and without the use of CAD and diagnostic accuracy numbers (true-positives, false-positives, true-negatives and false-negatives).

Study quality was assessed independently by the same two observers using the QUADAS tool [30, 31], disagreement was resolved by arbitration. This evidence-based tool is developed specifically to assess the quality of diagnostic accuracy studies and includes 14 quality items. The 14 items can be scored as “yes”, “no” or “unclear”. The total score can range from 0 to 14, in which 14 is the maximum attainable score.

### Statistical analysis

The performance of the radiologist in distinguishing breast lesions on MRI with and without the implementation of CAD was assessed. Besides the use of CAD, comparisons were made between radiologists with experience in imaging assessment and residents or radiologists with no or minor experience. Primary outcome was sensitivity and specificity at tumour level. Pooling of data was performed within the bivariate mixed-effects binary regression modelling framework. Model specification, estimation and prediction were carried out with xtmelogit in STATA. Using the model



**Fig. 1** Flow chart of search results, with reasons for exclusion and the total number of studies included

summary sensitivity and specificity were calculated, and a summary ROC curve was drawn (with AUC and confidence intervals). A forest plot was generated containing the individual study sensitivities and specificities with 95% confidence intervals (CI) and the pooled sensitivity and specificity estimates.

A test for heterogeneity was applied, using the  $I^2$  statistic [32]. This statistic calculates the percentage of total variation across studies that can be attributed to inter-study heterogeneity, ranging from 0 (no heterogeneity) to 100% (all variance due to heterogeneity). The presence of publication bias was visually assessed by producing a funnel plot. In STATA linear regression was performed of log odds ratios on the inverse root of effective sample sizes as a test for funnel plot asymmetry. The log odds ratios are defined as the log transformed diagnostic odds ratios, which are needed for the performance of linear regression.

Publication bias was considered present if there was a significant non-zero slope coefficient, ( $p < 0.10$ ), suggesting that only the small studies reporting a high sensitivity with CAD had been published, whereas the small studies reporting a lower sensitivity had not been published. Data were analysed in SPSS 16.0 (SPSS, Chicago, IL, USA), Meta Disc [33] and STATA SE version 11.0 (STATA, College Station, TX, USA).

## Results

### Study descriptives

The 10 studies included a total of 895 patients (range 29–329) with a total of 1264 breast lesions (range 33–469) of which 606 were classified as malignant (range 9–279) and 658 as benign (range 22–190) [20–29].

In 5 [23, 24, 26–28] studies a selection was made of patients with suspect findings based on mammography and ultrasound examinations. In the other 5 studies [20–22, 25, 29] patients with a suspect lesion on MRI were included. One of these 5 studies retrospectively searched the database of an ongoing MRI screening study of patients at high risk of breast cancer for BIRADS 3–5 lesions that were detected with MRI [22], and 2 studies included lesions that were not palpable and were not visible on mammography or ultrasound [20, 21]. In all 10 studies histology was used as the gold standard. In 4 studies a follow-up MRI after 6 or 24 months was performed [23, 25, 28, 29]; in the case of positive findings biopsy provided further histological assessment.

Mean study quality was 12.6, ranging from 10 to 14. Four studies were of maximum quality (Table 1) [20, 21, 26, 27].

### CAD systems

In all 10 studies the CAD systems (CADstream, DynaCAD, Fulltime point, 3-Time-Point Method and CAD-Gaea) incorporated precontrast medium (unenhanced) images and 2 (immediate and delayed) or all postcontrast medium (enhanced) images. The CAD systems compared pixel intensity values on the precontrast medium and immediated postcontrast medium series. If a pixel value increased above a user-specified minimum enhancement threshold, such as a 50 or 100% increase in enhancement, the pixel was regarded as meeting threshold enhancement. Once a pixel was identified as enhancing above the established threshold, the CAD systems compared pixel signal intensity values on the immediate and delayed postcontrast medium series to indicate washout enhancement, plateau enhancement or persistent enhancement. A specific colour or colour intensity was assigned to each pixel for different types of tissue enhancement. The end result of all CAD systems was a colour overlay on each MRI slice

**Table 1** Study characteristics of the 10 included studies (*SD* standard deviation, *NR* not reported, *P* prospective, *R* retrospective, *c* consecutive, *TB* tumour-based)

Study (author, ref., year of publication)	No. of patients	Study design	Quality score	Mean age (SD or range)	No. of lesions	No. of malignant	No. of benign	Type of analysis	MRI	CAD system
Arazi-Kleinman [22] 2009	53	R, c	13	47 (26–68)	56	22	34	TB	1.5T	CAD-Gaea
Meeuwis [27] 2009	65	R, c	14	49 (29–71)	71	49	22	TB	3.0T	CADstream
Baltzer [23] 2009	51	R, c	12	51 (13)	90	46	44	TB	1.5T	DynaCAD
Baltzer [24] 2009	329	P, c	13	53 (15–83)	469	279	190	TB	1.5T	DynaCAD
Veltman [29] 2009	NR	R,c	11	NR	52	25	27	TB	1.5T	3-Time-Point
Renz [28] 2008	48	P, c	11	51 (31)	88	43	45	TB	1.5T	DynaCAD Full-time Point
Hauth [25] 2008	137	R	10	NR	183	61	122	TB	1.5T	3-Time-Point
Williams [21] 2006	126	R, c	14	52 (27–86)	154	41	113	TB	1.5T	CADstream
Lehman [20] 2005	29	R, c	14	NR	33	9	24	TB	1.5T	CADstream
Kelcz [26] 2002	57	P, c	14	52 (31–80)	68	31	37	TB	1.5T	3-Time-Point

indicating regions of significant enhancement and providing details about enhancement type and extent.

#### CAD threshold enhancement

Six [20–24, 27] studies analysed the presence or absence of “threshold enhancement” at different minimum thresholds. Lehman et al. [20], Williams et al. [21] and Meeuwis et al. [27] used the CAD system CADstream. The sensitivity at the minimum thresholds of 50%, 80% and 100% remained

the same. The specificity increased at higher minimum thresholds. The study by Meeuwis et al. [27] showed a higher specificity than the other two studies. In the studies by Baltzer et al. [23, 24] DynaCAD was used. In these studies, the sensitivity decreased and the specificity increased at higher minimum thresholds. CAD-Gaea [22] had the same results as DynaCAD with respect to sensitivity and specificity, although CAD-Gaea had a lower level of specificity. Meeuwis et al. [27], using CADstream, reported the highest sensitivity and specificity (Table 2).

**Table 2** The sensitivity and specificity of a CAD system using the presence or absence of lesion enhancement at the user-specified minimum thresholds

Study	MRI	CAD System	No. of lesions	MRI assessed by using	Sensitivity	Specificity
Arazi-Kleinman [22]	1.5T	CAD-Gaea	56	Threshold 50%	100%	0%
				Threshold 80%	95.5%	14.7%
				Threshold 100%	72.7%	44.1%
Baltzer [23]	1.5T	DynaCAD	90	Threshold < 50%	100%	0%
				Threshold 50%–100%	84.8%	45.4%
				Threshold > 100%	52.1%	72.7%
Baltzer [24]	1.5T	DynaCAD	469	Threshold < 50%	100%	0%
				Threshold 50%–100%	86.4%	53.2%
				Threshold > 100%	52.0%	83.7%
Meeuwis [27]	3.0T	CADstream	71	Threshold 50%	97.9%	86.4%
				Threshold 100%	97.9%	90.9%
Williams [21]	1.5T	CADstream	154	Threshold 50%	92.7%	8.9%
				Threshold 100%	92.7%	23.0%
Lehman [20]	1.5T	CADstream	33	Threshold 50%	100%	25.0%
				Threshold 80%	100%	33.0%
				Threshold 100%	100%	50.0%

## Radiologist with or without CAD

In 8 out of 10 studies [22–29] the sensitivity and specificity of the radiologist or resident in assessing MR images with the use of CAD were measured (Table 3). The enhancement thresholds used were set up individually according to the radiologist's preference. Furthermore, in 4 [24, 27–29] out of these 8 studies the sensitivity and specificity of the radiologist or resident was also calculated without the use of CAD. In these four studies the radiologists or residents assessed the MR images as visual evaluation of contrast enhancement or by making a manual curve analysis by using the region of interest (ROI) method [24, 27–29]. The pooled sensitivity and specificity of radiologists and residents assessing breast lesions on MRI without the

implementation of CAD was 82% (95% CI: 72–90%) and 81% (95% CI: 74–87%), respectively. With CAD implementation they attained higher sensitivity scores (sensitivity: 89%, 95% CI: 83–93%; specificity: 81%, 95% CI: 76–85%) (Fig. 2). The sROC curve showed an AUC of 0.89 (Fig. 3). In 3 studies differentiation was made between radiologists with experience and residents with no or minimal experience [27–29]. The experience of those radiologists varied from 5 to 15 years (or > 500 MRIs). Residents had no more than 6 months (or < 50 MRIs) breast MRI experience. After stratification, the experienced radiologists showed a comparable pooled sensitivity of 89% with (95% CI: 81–94%) and without (sensitivity: 89%; 95% CI: 78–94%) CAD implementation. The pooled specificity of 86% (95% CI: 79–91%)

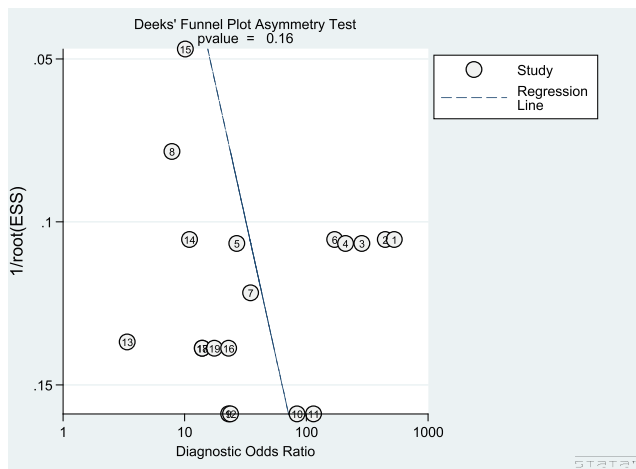
**Table 3** The performance of radiologists and residents in breast MRI diagnosis in terms of sensitivity and specificity with and without the use of a CAD system, specified for type of CAD and MRI system, number of lesions, and experience (*RAD* radiologist *RES* resident)

Study	MRI	CAD system	No. of lesions	MRI assessed by using	Experience	Sensitivity	Specificity
Arazi-Kleinman [22]	1.5T	CAD-Gaea	56	CAD+RAD	5 years	73.0%	56.0%
Meeuwis [27]	3.0T	CADstream	71	CAD+RAD1	> 5 years	88.5%	75.0%
			71	CAD+RAD2	> 5 years	92.3%	87.5%
			71	CAD+RES1	6 months	88.5%	93.8%
			71	CAD+RES2	0 months	84.6%	81.3%
			42	RAD (manual) <sup>a</sup>		84.6%	68.8%
Baltzer [23]	1.5T	DynaCAD	90	CAD+RAD	1–3 years	80.4%	72.7%
Baltzer [24]	1.5T	DynaCAD	469	CAD+RAD	>300 MRIs	78.8%	73.2%
			469	RAD (manual) <sup>a</sup>		75.3%	76.3%
			469	RAD (visual) <sup>b</sup>		72.4%	77.4%
Renz [28]	1.5T	DynaCAD	88	CAD+RAD1	> 500 MRIs	100%	86.7%
			88	CAD+RAD2	> 500 MRIs	95.3%	93.3%
			88	CAD+RAD3	< 50 MRIs	90.7%	73.3%
		Full-time point	88	CAD+RAD1		100%	84.4%
			88	CAD+RAD2		95.3%	91.1%
			88	CAD+RAD3		100%	66.7%
			88	RAD1 (visual) <sup>b</sup>		97.7%	84.4%
			88	RAD2 (visual) <sup>b</sup>		93.0%	93.3%
			88	RAD3 (visual) <sup>b</sup>		86.0%	77.8%
Veltman [29]	1.5T	3-Time-point	52	CAD+RES1	0 months	80%	78%
				CAD+RES2	3 months	80%	81%
				CAD+RES3	5 years	80%	85%
				CAD+RAD4	15 years	80%	78%
				RES1 (manual) <sup>a</sup>		68%	67%
				RES2 (manual) <sup>a</sup>		52%	81%
				RES3 (manual) <sup>a</sup>		72%	85%
				RAD4 (manual) <sup>a</sup>		84%	85%
Hauth [25]	1.5T	3-Time-point	183	CAD+RAD	3 years	60.7%	83.6%
Kelez [26]	1.5T	3-Time-point	68	CAD+RAD	>500 MRIs	87.0%	84.0%

<sup>a</sup> Manual: manual curve analysis by using the region of interest (ROI) method

<sup>b</sup> Visual: visual evaluation of contrast enhancement





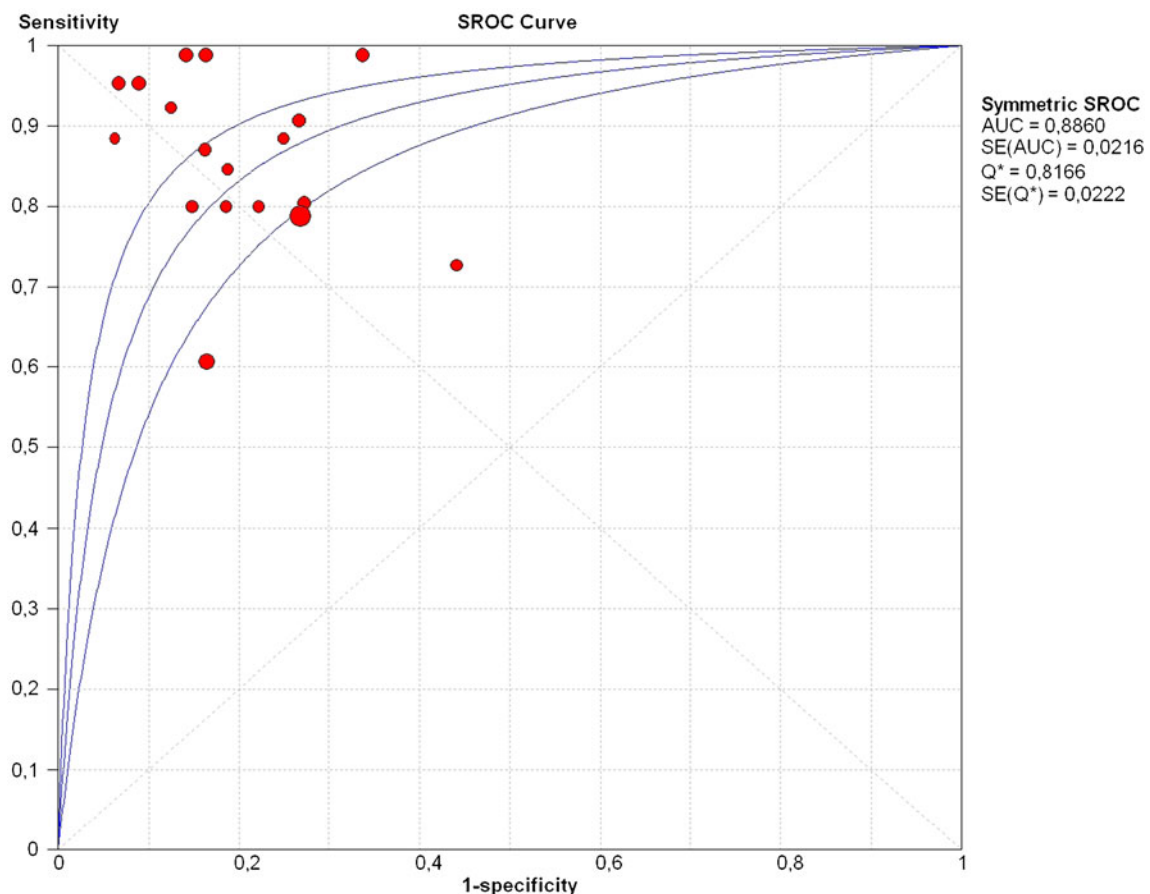
**Fig. 2** Funnel plot with log odds ratios on the inverse root of effective sample sizes for visualisation of publication bias

decreased to 82% (95% CI: 76–87%) with CAD. Residents or radiologists with less experience showed a pooled sensitivity of 72% (95% CI: 62–81%) and a pooled specificity of 79% (95% CI: 69–86%) when assessing breast lesions on MRI without CAD. With the use of

CAD, their sensitivity increased to 89% (95% CI: 80–94%), whereas their specificity remained comparable (specificity: 78%; 95% CI: 69–84%) (Table 4).

#### Analyses of heterogeneity

Moderate to substantial heterogeneity was observed among the eight studies exploring the sensitivity and specificity of radiologists assessing MR images with and without the implementation of CAD (sensitivity: no CAD:  $I^2$ : 78%,  $p < 0.0001$ , with CAD:  $I^2$ : 80%,  $p < 0.0001$ ; specificity: no CAD:  $I^2$ : 46%,  $p = 0.007$ , with CAD:  $I^2$ : 55%,  $p = 0.002$ ). After stratification of radiologists with experience and residents with no or minimal experience, heterogeneity did not change for experienced radiologists (sensitivity: no CAD:  $I^2$ : 79%,  $p < 0.0001$ , with CAD:  $I^2$ : 83%,  $p < 0.0001$ ), and residents (sensitivity: no CAD:  $I^2$ : 79%,  $p = 0.009$ ; with CAD:  $I^2$ : 64%,  $p = 0.02$ ). Concerning specificity, heterogeneity dropped to low to moderate (radiologists without CAD: specificity:  $I^2$ : 56%,  $p = 0.04$ , radiologists with CAD: specificity:  $I^2$ : 62%,  $p = 0.01$ ; residents without CAD: specificity:  $I^2$ : 33%,  $p = 0.22$ ; residents with CAD: specificity:  $I^2$ : 24%,  $p = 0.26$ ). Because of heterogeneity and possible



**Fig. 3** Summary ROC curve regarding the studies of radiologists and residents using a CAD system

**Table 4** Results of pooled sensitivity and specificity (95% CI) of the radiologist in assessing breast lesions on MRI with and without the use of a CAD system in general, stratified for experienced radiologists and residents with no or less experience (RANDOM effects model)

Outcome or subgroup	Studies <sup>a</sup>	Sensitivity (95% CI)	Specificity (95% CI)
Radiologist no CAD, general	4 <sup>b</sup>	82% (72%–90%)	81% (74%–87%)
Radiologist with CAD, general	8 <sup>c</sup>	89% (83%–93%)	81% (76%–85%)
Experienced radiologist no CAD	4 <sup>b</sup>	89% (78%–94%)	86% (79%–91%)
Experienced radiologist with CAD	8 <sup>c</sup>	89% (81%–94%)	82% (76%–87%)
Resident no CAD	3 <sup>d</sup>	72% (62%–81%)	79% (69%–86%)
Resident with CAD	3 <sup>d</sup>	89% (80%–94%)	78% (69%–84%)

<sup>a</sup> In studies in which more than one radiologist/resident (blinded) assessed the images, the pooled calculation was based on all relevant radiologists in that study

<sup>b</sup> Meeuwis [27], Baltzer [24], Renz [28], Veltman [29]

<sup>c</sup> Arazi-Kleinmann [22] Meeuwis [27], Baltzer [23], Baltzer [24], Renz [28], Veltman [29], Hauth [25], Kelcz [26]

<sup>d</sup> Meeuwis [27], Renz [28], Veltman [29]

unmeasured variance at the study level a random-effects model was used to obtain all pooled estimates, as this model interprets the available data with more caution and uses broad confidence intervals.

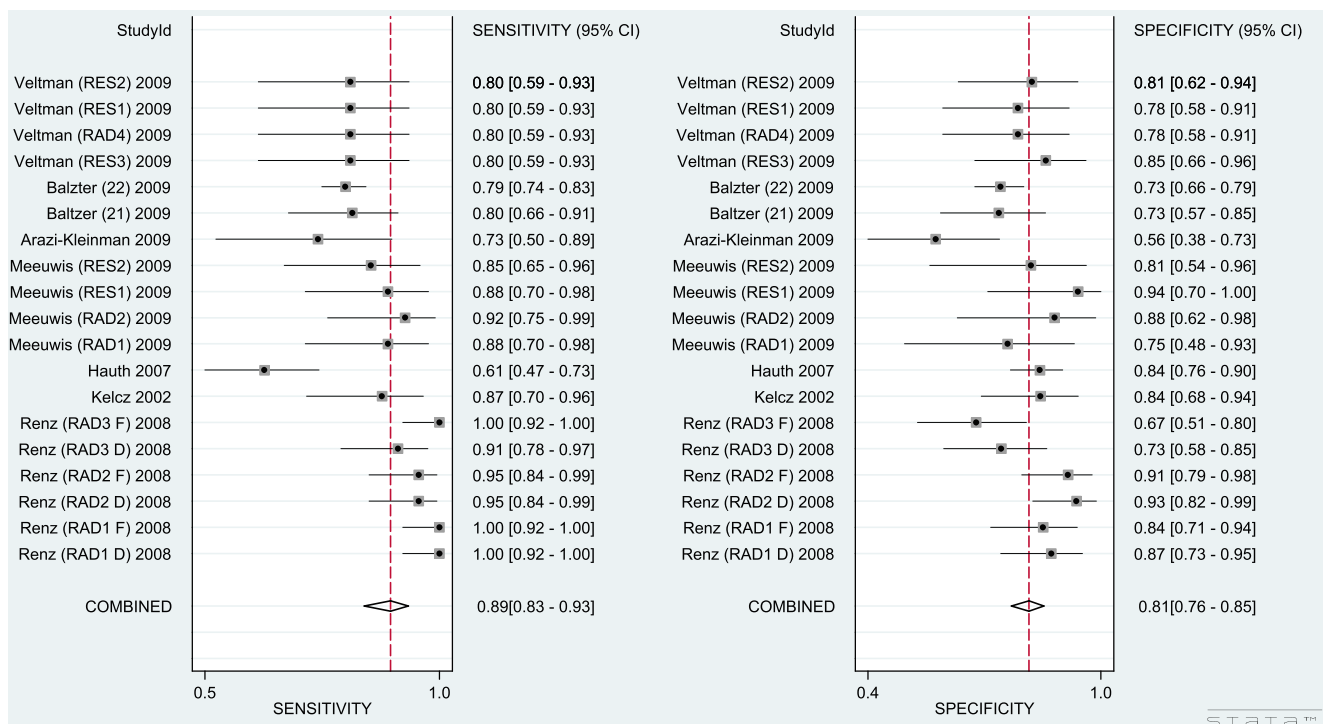
#### Assessment publication bias

A non-significant non-zero slope coefficient (p-value=0.16) indicated that there was no evidence of publication bias (Fig. 4). This suggests that we most likely did not miss studies with a negative outcome.

#### Discussion

This meta-analysis shows that the pooled sensitivity and specificity of the experienced radiologist for the assessment of breast lesions with MRI remains comparable with the implementation of CAD. Residents or radiologists with less experience seem to attain a higher sensitivity with CAD implementation, although not significant.

All selected studies were of high quality, so it is likely that the quality of the studies did not have a significant impact on the results of the meta-analysis. Furthermore, there was no

**Fig. 4** Forest plot of pooled sensitivity and specificity of radiologists and residents assessing breast lesions on MRI with the use of a CAD system



evidence of publication bias and therefore it is not expected that the meta-analysis overestimates the effect of CAD evaluation.

The ten studies used different indications for breast MRI and there was a wide variation in the number and tissue type of lesions selected [20–29]. This resulted in the greater heterogeneity. Therefore, we used a random-effects model that interprets the results with more caution. Furthermore, there was an indication of selection bias. In all studies the radiologists only assessed MRIs with lesions ( $\geq$  BIRADS 2) and discriminated between benign and malignant. This selection increased the prevalence of breast malignancy in the study population compared with the target population. The lesion selection could have influenced the performance of the radiologist.

Six studies analysed the influence of the presence or absence of “threshold enhancement” at different minimum thresholds. Of those six, the study by Meeuwis et al. [27] resulted in the highest sensitivity and specificity. This result could be explained by the fact that a 3.0T MRI system was applied, which has a better performance than the 1.5T MRI scanner which was used in the other studies.

With CADstream the sensitivity at a higher enhancement threshold remained the same, i.e. the same malignant lesions enhanced at the 50%, 80% and 100% thresholds [20]. The remaining false-negative enhancing malignant lesions showed no enhancement with CADstream due to a noise filtering process leading to failure of automatic analysis of small areas of enhancement [21, 27]. The specificity of CADstream increased at higher enhancement thresholds, which means that at a higher threshold benign lesions did not enhance [20]. Therefore, absence of lesion enhancement at higher thresholds helps to improve the discrimination between benign and malignant lesions. In comparison to the study by Meeuwis et al. [27], the low specificity of the study by Williams et al. [21] is most likely due to the high prevalence ( $n=22/71$  versus  $n=113/154$ ) and the large tissue type variation of benign lesions.

With the DynaCAD software the specificity performance was analogous to that of the CADstream software. The sensitivity of DynaCAD however, decreased at higher threshold enhancements, not visualising all malignant enhancements, resulting in false-negative lesions [23].

Residents or radiologists with no or less experience achieved a higher sensitivity when they were accompanied by a CAD system for discrimination between breast lesions on MRI. The change in sensitivity after using CAD was not significant. Nevertheless, a considerable increase could be seen (sensitivity from 72%; 95% CI: 62–81% to 89%; 95% CI: 80–94%). This increase could be a result of the fact that CAD brings more enhancing lesions to the attention of the resident or inexperienced radiologist. Therefore, it seems

that they benefit from CAD when assessing breast lesions with MRI. However, more research must be conducted to verify these results.

The performance of the experienced radiologists showed a non-significant decrease in specificity from 86% (95% CI: 79–91%) without CAD to 82% (95% CI: 76–87%) with CAD. A clarification for this observation could be that CAD systems are only based on the enhancement dynamic, without regarding the morphology of the lesion. As a consequence, the use of CAD could lead to a higher number of enhancing lesions, part of which could be assigned as benign on the basis of morphology. The experienced radiologists can be misled by the enhancement pattern of CAD, resulting in a decrease in specificity. Therefore, it is important that experienced radiologists are aware of this.

In conclusion, concerning the assessment of MR images CAD has little influence on the sensitivity and specificity of the performance of radiologists experienced in breast MRI diagnosis. Therefore, breast MRI interpretation by radiologists remains essential. Residents or radiologists with less experience seem to benefit from a CAD system when performing breast MRI evaluation.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. DeMartini W, Lehman C (2008) A review of current evidence-based clinical applications for breast magnetic resonance imaging. *Top Magn Reson Imaging* 19:143–150
2. DeMartini W, Lehman C, Partridge S (2008) Breast MRI for cancer detection and characterization: a review of evidence-based clinical applications. *Acad Radiol* 15:408–416
3. Kuhl CK (2007) Current status of breast MR imaging. Part 2. Clinical applications. *Radiology* 244:672–691
4. Mann RM, Kuhl CK, Kinkel K, Boetes C (2008) Breast MRI: guidelines from the European Society of Breast Imaging. *Eur Radiol* 18:1307–1318
5. Orel S (2008) Who should have breast magnetic resonance imaging evaluation? *J Clin Oncol* 26:703–711
6. Orel SG, Schnall MD (2001) MR imaging of the breast for the detection, diagnosis, and staging of breast cancer. *Radiology* 220:13–30
7. Macura KJ, Ouwerkerk R, Jacobs MA, Bluemke DA (2006) Patterns of enhancement on breast MR images: interpretation and imaging pitfalls. *Radiographics* 26:1719–1734
8. Schnall MD, Blume J, Bluemke DA, DeAngelis GA, DeBruhl N, Harms S, Heywang-Kobrunner SH, Hylton N, Kuhl CK, Pisano ED, Causer P, Schnitt SJ, Thickman D, Stelling CB, Weatherall PT, Lehman C, Gatsonis CA (2006) Diagnostic architectural and dynamic features at breast MR imaging: multicenter study. *Radiology* 238:42–53

9. Szabo BK, Aspelin P, Wiberg MK, Bone B (2003) Dynamic MR imaging of the breast. Analysis of kinetic and morphologic diagnostic criteria. *Acta Radiol* 44:379–386
10. Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, Heywang-Kobrunner SH, Hylton N, Kuhl CK, Lehman C, Pisano ED, Causer P, Schnitt SJ, Smazal SF, Stelling CB, Weatherall PT, Schnall MD (2004) Magnetic resonance imaging of the breast prior to biopsy. *JAMA* 292:2735–2742
11. Lee CH (2004) Problem solving MR imaging of the breast. *Radiol Clin North Am* 42:919–934
12. Peters NH, Borel Rinkes IH, Zuithoff NP, Mali WP, Moons KG, Peeters PH (2008) Meta-analysis of MR imaging in the diagnosis of breast lesions. *Radiology* 246:116–124
13. Kuhl CK, Schmützler RK, Leutner CC, Kempe A, Wardelmann E, Hocke A, Maringa M, Pfeifer U, Krebs D, Schild HH (2000) Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: preliminary results. *Radiology* 215:267–279
14. Moy L, Elias K, Patel V, Lee J, Babb JS, Toth HK, Mercado CL (2009) Is breast MRI helpful in the evaluation of inconclusive mammographic findings? *Am J Roentgenol* 193:986–993
15. Vassiou K, Kanavou T, Vlychou M, Poultsidi A, Athanasiou E, Arvanitis DL, Fezoulidis IV (2009) Characterization of breast lesions with CE-MR multimodal morphological and kinetic analysis: comparison with conventional mammography and high-resolution ultrasound. *Eur J Radiol* 70:69–76
16. Hsung JM, Sonnad SS, Schwartz JS, Langlotz CP (1999) Accuracy of MR imaging in the work-up of suspicious breast lesions: a diagnostic meta-analysis. *Acad Radiol* 6:387–397
17. Rothenberg RM (2006) Computer-aided detection of malignancy with magnetic resonance imaging of the breast. *Technol Eval Cent Asses Program Exec Summ* 21:1–3.
18. Wood C (2005) Computer Aided Detection (CAD) for breast MRI. *Technol Cancer Res Treat* 4:49–53
19. Castellino RA (2005) Computer aided detection (CAD): an overview. *Cancer Imaging* 5:17–19
20. Lehman CD, Peacock S, DeMartini WB, Chen X (2006) A new automated software system to evaluate breast MR examinations: improved specificity without decreased sensitivity. *Am J Roentgenol* 187:51–56
21. Williams TC, DeMartini WB, Partridge SC, Peacock S, Lehman CD (2007) Breast MR imaging: computer-aided evaluation program for discriminating benign from malignant lesions. *Radiology* 244:94–103
22. Arazi-Kleinman T, Causer PA, Jong RA, Hill K, Warner E (2009) Can breast MRI computer-aided detection (CAD) improve radiologist accuracy for lesions detected at MRI screening and recommended for biopsy in a high-risk population? *Clin Radiol* 64:1166–1174
23. Baltzer PA, Renz DM, Kullnig PE, Gajda M, Camara O, Kaiser WA (2009) Application of computer-aided diagnosis (CAD) in MR-mammography (MRM): do we really need whole lesion time curve distribution analysis? *Acad Radiol* 16:435–442
24. Baltzer PA, Freiberg C, Beger S, Vag T, Dietzel M, Herzog AB, Gajda M, Camara O, Kaiser WA (2009) Clinical MR-mammography: are computer-assisted methods superior to visual or manual measurements for curve type analysis? A systematic approach. *Acad Radiol* 16:1070–1076
25. Hauth EA, Jaeger H, Maderwald S, Muhler A, Kimmig R, Forsting M (2008) Quantitative parametric analysis of contrast-enhanced lesions in dynamic MR mammography. *Radiologe* 48:593–600
26. Kelcz F, Furman-Haran E, Grobeld D, Degani H (2002) Clinical testing of high-spatial-resolution parametric contrast-enhanced MR imaging of the breast. *Am J Roentgenol* 179:1485–1492
27. Meeuwis C, van de Ven SM, Stapper G, Fernandez Gallardo AM, van den Bosch MA, Mali WP, Veldhuis WB (2009) Computer-aided detection (CAD) for breast MRI: evaluation of efficacy at 3.0 T. *Eur Radiol* 20:522–528
28. Renz DM, Baltzer PA, Kullnig PE, Bottcher J, Vag T, Gajda M, Camara O, Runnebaum IB, Kaiser WA (2008) Clinical value of computer-assisted analysis in MR mammography. A comparison between two systems and three observers with different levels of experience. *Rofo* 180:968–976
29. Veltman J, Mann RM, Meijer FJ, Heesakkers RA, Heufke M, Blickman JG, Boetes C (2010) The additional value of three time point color coding in dynamic contrast-enhanced MRI of the breast for inexperienced and experienced readers. *Eur J Radiol* 74:514–518
30. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 3:25
31. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J (2006) Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 6:9
32. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *BMJ* 327:557–560
33. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 6:31